APPENDIX 1 – Clean copies of replacement paragraphs

Replace the paragraphs at p.15, lines 21 to 33 with the following replacement paragraphs.

Figure 4 shows the 3414 base sequence of the coding region of *Naegleria* gene TTK(SEQ ID NO: 1) ; the first segment of this gene (underlined) encodes thiaminase I.

Figure 5 shows the amino acid sequence encoded by *Naegleria* gene TTK(SEQ ID NO: 2).

Figure 6 shows the DNA sequence of the 1068 base segment that encodes the Naegleria thiaminase I(SEQ ID NO: 3), as obtained from *Naegleria* gene TTK. This segment, expressed as pNB1+, encodes catalytically active thiaminase.

B1

Figure 7 shows the 356 amino acid sequence encoded by the *Naegleria* gene segment expressed in pNB1+, along with the DNA codons(SEQ ID NOS 3 and 4).

Figure 8 is an alignment comparing the amino acid sequence of *Naegleria* thiaminase I to other homologous sequences, specifically the thiaminase I of *Bacillus* thiaminase I and segments of several transketolases, which we found show limited homology to the encoded sequences of the two sequenced thiaminase I proteins(SEQ ID NOS 5-11, 4 and 12-14, respectively, as they occur).

Clean copies of replacement paragraphs to replace the paragraphs at p.24, line 23 to p.26, line 6.

3-8. Genomic clone isolation. Three peptide fragment sequences derived from analysis of the gel purified putative thiaminase were obtained. Peptides A (SEQ ID NO: 12) (ASDLPQSGDQVNK), B(SEQ ID NO: 13) (TILDSTVVASQR), and C(SEQ ID NO: 14)

B2

(SSNFYAQLSQQFDAK) were compared to the peptide sequence of *Bacillus subtilis* thiaminase I precursor to determine the best estimated local alignment for each fragment using LALIGN (FASTA v.2.0u54 (X. Huang and W. Miller (1991) Adv. Appl. Math. 12:337-357)). Sets of forward and reverse degenerate oligonucleotide primers were designed based on these alignments, and synthesized using an Expedite™ model 899 DNA synthesizer using phosphoramidite synthesis chemistry (PerSeptive Biosystems, Framingham, MA). Peptide A forward primer AF1(SEQ ID NO: 15) (5'-CARWSIGGHGAYCARG-3') and peptide C reverse primer CR1(SEQ ID NO: 16) (5'-TTIGCRTCRAAYTGYTG) generated an amplimer of ~400 bp as expected based on the *B. subtilis* alignments, was designated AC400 and chosen for gel purification using a QIAEXII gel purification kit (QIAGEN, Santa Clarita, CA). Fluorescent DNA sequencing reactions using the CR1 degenerate oligonucleotide and the gel purified AC400 PCR DNA fragment were performed using an Applied Biosystems 373 Stretch DNA Sequencer, and ABI PRISM™ dye terminator cycle sequencing reagents using DNA polymerase FS following manufacturer recommendations (PE Biosystems, Foster City, CA). The resulting sequence facilitated the design of two homologous internal oligonucleotides a forward primer AC1(SEQ ID NO: 17) (5'-TGTCGGATATAGTGAAAGTATG-3') and a reverse primer AC2(SEQ ID NO: 18) (5'-AACCTTTTGCTTTTCATCACAC-3'), which were used to complete the 431 bp DNA sequence of AC400 that encodes a peptide with strong homology to *B. subtilis* thiaminase I precursor. AC400 was labeled using the random hexamer primer method and used to screen a *Naegleria gruberi* strain NEG EMBL3 genomic DNA library ((Lai et al., 1988b)), with each of the positive clones sequenced with oligonucleotides AC1 and AC2. The resulting genomic clone TTK containing all of the 431 bp of the AC400 hybridization probe was chosen for more extensive DNA sequence analysis using oligonucleotide primers.

3-9. Sequence analysis and expression cassette construction. Analysis of 3414 bp of genomic sequence from TTK defined an ORF of 3078 bp (Fig. 4) encoding a 1025 amino acid peptide of MW 112,141 daltons (Fig. 5), with the amino terminal domain amino acid residues Met1 to Leu356 encoding a peptide of MW 39,079 daltons having strong homology to *B. subtilis* thiaminase I precursor. This homology is shown is Fig. 8, as well as the positions of the three

peptide sequences from *Naegleria* thiaminase that were used to find the gene. This region was chosen for expression in pET-22b(+), with the expression cartridge generated by PCR amplification of TTK target DNA using *pfu* polymerase (STRATAGENE, La Jolla, CA). An NdeI site was introduced at the ATG start codon using a forward primer CDA1(SEQ ID NO: 19) (5'-GAGATATA**CATATG**TCCACTCAACCAAAGAC-3'), and a TAA stop codon followed by a BamHI site was introduced after Leu356 using the reverse primer CDA2(SEQ ID NO: 20) (5'-TAT**GGATCC**TTAAAGGAATGGTCTCAAGACACC-3'). The resulting 1091 bp amplimer was digested with NdeI and BamHI, gel purified and the expression cassette ligated into pET-22b(+) followed by transformation into E. coli XL1-blue™ (STRATAGENE, La Jolla, CA). Sequencing of the resulting plasmid pNB1+ verified the expected 1068 bp ORF (Fig. 6) including the potential catalytic Cys118, as well as the junctions essential for expression of the 356 amino acid Met1–Leu356 peptide (Fig. 7).

B2

Clean copy of replacement paragraph to replace the paragraph at p.26, line 23 to p.27, line 6.

3-11. PCR site-directed mutagenesis. PCR site directed mutagenesis of the potentially catalytic cys118 codon (TGC) to ser118 (AGC) was performed using a variation on the method of Barettino et al. (1994) N.A.R. 22(3):541-542. The position of the catalytic Cys codon is shown in Fig. 8. A mutagenic coding strand (reverse) primer CDA3 (SEQ ID NO: 21) (5'-CAA TAA AAA GTT TGA **GCT** CAA GTA TTG-3') was used in conjunction with the forward primer CDA1 to amplify a 380 bp mutagenic megaprimer. Following gel purification the megaprimer was extended using the reverse primer CDA2 in a linear PCR extension reaction. The resulting 1091 bp fragment, containing a GCT (encoding Ser) instead of GCA (encoding Cys118) was digested with NdeI and BamHI, subcloned, sequenced and expressed as detailed above. The resulting mutant clone, named pNB1-S , was confirmed to have a Cys118Ser substitution, and it

B3

B3

tested negative for enzymatic activity on agar, in the spectrophotometer, and by a cell death assay. The result of this experiment is listed in Table 2.

Clean copies of replacement paragraphs to replace the paragraphs at p.29, line 27 to p.30, line 2.

3) Local similarity to active site regions as determined by BLAST. For example, the catalytic domain segment of 13 residues surrounding the active site Cys, (SEQ ID NO: 22) VYGFPQYLCSNFL, would be expected to give an identity of 8 of the 13 amino acids (see Fig. 8).

B4

Once candidate thiaminases are recognized by these criteria, they can be expected to show other features. One example is the six amino acid sequence (SEQ ID NO: 23) GYSESM that starts at residue 228 of Naegleria thiaminase I (Fig. 7). This is part of the pyrimidine coordinate residues, and one would expect a match of ≥5 of the 6 residues.